Optimal Control of Heteroscedastic Macroeconomic Models

Vito Polito[a*] and Peter Spencer[b]

[a]University of Bath, United Kingdom

[b]University of York, United Kingdom

May 2015

Abstract

This paper analyses the implications of heteroscedasticity for optimal macroeconomic policy and welfare. We find that changes in the variance structure driven by exogenous processes like GARCH affect welfare but not the optimal feedback rule. However, changes in the variance structure driven by state-dependent processes affect both. We also derive Certainty-Equivalent Transformations of state-dependent volatility models that allow standard quadratic dynamic programming algorithms to be employed to study optimal policy. These results are illustrated numerically using a reduced-form model of the U.S. economy in which changes in volatility are driven by a GARCH process and the rate of inflation

Keywords: Heteroscedasticity, Optimal Control, Macroeconomic Volatility, Optimal Monetary Policy.

JEL classification: C32, C61, E52

1 Introduction

This paper explores the implications of heteroscedastic disturbances for the analysis of optimal policy. Our analysis is based on the observation that most time-varying volatility models are essentially quadratic and therefore fit nicely into the linear-quadratic framework of the optimal linear regulator problem, allowing a rigorous analysis of their policy implications.

*Correspondence to: Vito Polito, Department of Economics, 3 East, University of Bath, Bath, BA2 7AY, UK. E-mail: v.polito@bath.ac.uk

Three key findings emerge from this research. First, the certainty-equivalence (CE) principle still holds in macroeconomic models with heteroscedastic disturbances if changes in the variance structure are exogenous. Examples of this type include ARCH, GARCH or stochastic volatility processes that, like homoscedastic volatility, influence the welfare loss but not the optimal policy.

Second, we find that when changes in the variance structure are related to the variables describing the state of the economy, the CE principle no longer holds and the specification of the variance structure does influence the optimal decision rule. State-dependent models of the variance structure are extensively used in the literature on the term structure of interest rates and have been employed more recently in macroeconomic models. They allow the variance of the shocks to depend in both a linear and a quadratic way upon the state variables. The classic example is the Cox, Ingersoll and Ross (1985) square root volatility model, in which the variance depends linearly upon the nominal rate of interest. Dothan (1978) and Courtadon (1982) develop models in which the variance is a quadratic function of the interest rate. Engle (1982) uses lagged values of the regressors as a way of generalizing ARCH variance specifications. Recent examples of macro-finance models with a state-dependent variance component include Spencer (2008), Bekaert, Cho and Moreno (2010) and Campbell et al. (2014). We show that there is a hierarchy of effects if the error structure is heteroscedastic. Quadratic state dependence affects all of the coefficients in the optimal feedback rule, as well as welfare. Linear state dependence affects the intercept of the optimal feedback rule and welfare. GARCH reinforces these state-dependent effects. However, on its own, GARCH only affects welfare.

Third, we derive a Certainty-Equivalent Transformation (CET) of the heteroscedastic optimal linear regulator problem with state-dependent volatility. This uses change of variable techniques to write the problem as in the canonical homoscedastic form. This allows researchers to use standard optimal control techniques to analyze optimal policy rules and welfare losses. The transformation shows that state dependent volatility changes the effective welfare cost of variables like inflation and interest rates that may influence volatility. CETs of dynamic optimization problems are common in both finance and macroeconomics. Hansen and Sargent (2008) provide a textbook description of these in the context for robust optimal control of models with dynamic misspecification. Our work complements theirs by examining the effect of stochastic misspecification.

We illustrate the theoretical results numerically using a small-scale VAR model of the U.S. economy as a laboratory to revisit one of the most popular applications of dynamic programming in macroeconomics: the analysis of optimal monetary policy. The model is estimated by Maximum Likelihood (ML). The variance structure of the VAR includes both GARCH and inflation-dependent components that are highly significant statistically. This specification is consistent with the Okun-Friedman-Ball hypothesis that macroeconomic uncertainty is related to the rate of inflation, see Okun (1971), Friedman (1977) and Ball (1992). Fountas, Karanasos and Kim (2002) and Caporale and Kontonikas (2009) argue that an increase in inflation should lead to a monetary tightening response to limit the increase in macroeconomic volatility. Our model formalizes this proposition and provides an argument for a low inflation target as well as a more aggressive response to inflation shocks.[1] Our numerical results show how misspecification of the variance structure can lead researchers to mismeasuring both the welfare cost of inflation and the potential gains from optimization.

The rest of the paper proceeds as follows. Section 2, supported by Appendices A and B, sets out the general solution of the optimal linear regulator problem with heteroscedastic disturbances; shows how this depends upon the source of heteroscedasticity; and derives the CET under state dependence. Section 3, supported by Appendices B, C and D, describes the empirical application used to illustrate the theoretical results. Section 4 concludes by summarizing the findings of this research. Appendix F suggests extensions of the model framework and highlights avenues for future research.[2]

2 Optimal control of heteroscedastic macroeconomic models

This section presents a general framework for optimization problems with linear-quadratic heteroscedasticity, based on the Bellman equation. This includes expectations of quadratic

[1] There are arguments that point in the opposite direction, suggesting a higher target (Blanchard *et al* (2010)) and a less aggressive response (Sack (2000)).

[2] Appendices are available online.

forms in the state variables, which involve both means and variances. Since the latter are linear-quadratic functions of the state variables, the value function remains quadratic and the decision rules linear.

2.1 Specification

Let $\beta \in (0,1)$ be a discount factor and E_t denoting mathematical expectation conditional on information available in period t. Consider a decision maker that wants to choose an infinite sequence of controls $\{\mathbf{i}_t\}_{t=0}^{\infty}$ to minimize the quadratic loss function

$$V_t = \sum_{t=0}^{\infty} \beta^t E_t \begin{bmatrix} (\mathbf{x}_t - \mathbf{x}^*)' \mathbf{R} (\mathbf{x}_t - \mathbf{x}^*) + (\mathbf{i}_t - \mathbf{i}^*)' \mathbf{W} (\mathbf{i}_t - \mathbf{i}^*) \\ +2 (\mathbf{x}_t - \mathbf{x}^*)' \mathbf{H} (\mathbf{i}_t - \mathbf{i}^*) \end{bmatrix} \tag{1}$$

subject to the first-order stochastic linear difference equation

$$\mathbf{x}_{t+1} = \mathbf{A}\mathbf{x}_t + \mathbf{B}\mathbf{i}_t + \mathbf{w}_{t+1} \tag{2}$$

with $\mathbf{x}_0$ given. In the above, $\mathbf{x}_t$ is a $n \times 1$ vector of state or non-policy variables; $\mathbf{i}_t$ is a $q \times 1$ vector of control or policy variables; $\mathbf{R}$ is a $n \times n$ positive definite symmetric matrix; $\mathbf{W}$ is a $q \times q$ non-negative definite symmetric matrix; $\mathbf{H}$ is a $n \times q$ matrix; $\mathbf{x}^*$ and $\mathbf{i}^*$ are vectors of targets of dimension n and q respectively; $\mathbf{A}$ is a $n \times n$ matrix of coefficients; $\mathbf{B}$ is a $n \times q$ matrix of coefficients; and $\mathbf{w}_{t+1}$ is a $n \times 1$ vector of independently and identically distributed (i.i.d.) random variables with mean vector zero and heteroscedastic covariance matrix

$$\mathbf{\Sigma}_{t+1} = \mathbf{K} + \mathbf{C}'\mathbf{w}_t\mathbf{w}_t'\mathbf{C} + \mathbf{G}'\mathbf{\Sigma}_t\mathbf{G} + \mathbf{L}\mathbf{x}_t'\mathbf{s} + \mathbf{Q}\mathbf{x}_t'\mathbf{S}\mathbf{x}_t, \tag{3}$$

with $\mathbf{w}_0$ and $\mathbf{\Sigma}_0$ given. The five terms on the right side of (3) denote the homoscedastic, ARCH, GARCH, linear and quadratic state-dependent components of the covariance matrix respectively: $\mathbf{K}$ is a positive definite $n \times n$ matrix; $\mathbf{C}$, $\mathbf{G}$, $\mathbf{L}$ and $\mathbf{Q}$ are $n \times n$ matrices that are not necessarily symmetric.[3] The vector $\mathbf{s}$ and the matrix $\mathbf{S} = \mathbf{s}\mathbf{s}'$ select the variable(s)

[3] We assume that the sufficient conditions for the stability of the solution to the linear regulator problem are met, namely (i) the matrices $\overline{\mathbf{B}} = \mathbf{B} - \mathbf{C}\mathbf{R}^{-1}\mathbf{H}'$ and $\mathbf{C}$ are *stabilizable* and (ii) the matrices $\overline{\mathbf{\Lambda}} = \mathbf{\Lambda} - \mathbf{H}\mathbf{R}^{-1}\mathbf{H}'$ and $\mathbf{R}$ are positive semidefinite. See Ljungqvist and Sargent (2004), Appendix B.3, which transforms the system by removing the off diagonal $\mathbf{H}$ terms to get $\bar{\mathbf{B}}$ and $\bar{\mathbf{\Lambda}}$, and pages 116-118

entering the linear and quadratic components of the covariance matrix respectively.

A wide range of macro models can be written in the form of the first-order stochastic linear difference equation (2). For example, equation (2) can describe the non-policy part of VAR models such as those used for the measurement of macroeconomic shocks by Bernanke and Mihov (1998); and for optimal control by Sack (2000) and Polito and Wickens (2012). It encompasses the Rudebusch and Svensson's (1999) central bank model, which has been extensively employed for the analysis of U.S. monetary policy. Equation (2) is also consistent with the solution of a linear rational expectations model, as in Blanchard and Kahn (1980) for example.

Further, the specification of the covariance matrix in equation (3) includes classes of time-varying volatility models widely employed in macroeconomics and finance. For example, it encompasses Engel (1982)'s model, which exhibits ARCH and inflation-conditional dependence; and the constant-covariance version of the BEKK model of Engle and Kroner (1995). The linear state-dependent component of the covariance matrix encompasses models in which time-varying volatility is related to a state variable. Prominent examples of this include, the Cox, Ingersoll and Ross (1985) model where volatility is linked to the nominal rate of interest; the inflation-conditional volatility model postulated by the Okun-Friedman-Ball hypothesis and the macro-finance model formulated by Campbell et al. (2014) in which volatility is linked to the output gap. Dothan (1978) and Courtadon (1982) give examples of quadratic state-dependent volatility models where the driving factor is the nominal rate of interest.

2.2 General solution

To find the policy function we need to express the optimal value of the original problem given arbitrary initial conditions. In a standard homoscedastic quadratic dynamic programming problem, the value function includes a constant term (for the steady-state variance) and the state vector $\mathbf{x}_0$ (for any initial disequilibrium). This can also allow for linear-quadratic

which then applies the stability conditions. These assumptions imply that the solutions to the homoscedastic and the GARCH optimal linear regulator problems - which are both certainty equivalent - are stable. In section 2.2.4, we use the CET to infer that the stability properties of the optimal solution are preserved under state-dependent volatility.

(LQ) terms in the dynamic specification of the variance structure. The presence of ARCH and GARCH terms in (3) implies that the value function depends also on the initial values for $\mathbf{w}_0$ and $\mathbf{\Sigma}_0$. Thus we try a value function of the form:

$$V = V\left(\mathbf{x}, \mathbf{w}, \mathbf{\Sigma}\right) = k - 2\mathbf{x}'\mathbf{p} + \mathbf{x}'\mathbf{P}\mathbf{x} + \mathbf{c}'\mathbf{w}\mathbf{w}'\mathbf{c} + \mathbf{g}'\mathbf{\Sigma}\mathbf{g}, \tag{4}$$

where k is a scalar; $\mathbf{P}$ is a $n \times n$ positive semidefinite symmetric matrix, $\mathbf{p}$, $\mathbf{c}$ and $\mathbf{g}$ are $n \times 1$ vectors.[4] After using the transition law (2) and the covariance matrix (3) to eliminate next period's states in (1), taking expectations and recognizing that the vector $\mathbf{w}$ is orthogonal to $\mathbf{x}$ and $\mathbf{i}$, the Bellman equation becomes:

$$\begin{aligned}
&\overbrace{V = \min_{\mathbf{i}}\{\left(\mathbf{x}-\mathbf{x}^*\right)'\mathbf{R}\left(\mathbf{x}-\mathbf{x}^*\right) + \left(\mathbf{i}-\mathbf{i}^*\right)'\mathbf{W}\left(\mathbf{i}-\mathbf{i}^*\right) + 2\left(\mathbf{x}-\mathbf{x}^*\right)'\mathbf{H}(\mathbf{i}-\mathbf{i}^*) + \beta k + I}^{\text{Homoscedastic Bellman equation}} \\
&+ \underbrace{\beta\left[tr(\mathbf{P}\mathbf{C}'\mathbf{w}\mathbf{w}'\mathbf{C}) + \mathbf{c}'(\mathbf{K} + \mathbf{C}'\mathbf{w}\mathbf{w}'\mathbf{C})\mathbf{c}\right]}_{\text{ARCH effect}} \\
&+ \underbrace{\beta\left[tr\left(\mathbf{P}\mathbf{G}'\mathbf{\Sigma}\mathbf{G}\right) + \mathbf{g}'(\mathbf{K} + \mathbf{C}'\mathbf{w}\mathbf{w}'\mathbf{C} + \mathbf{G}'\mathbf{\Sigma}\mathbf{G})\mathbf{g} + \mathbf{c}'\mathbf{G}'\mathbf{\Sigma}\mathbf{G}\mathbf{c}\right]}_{\text{Additional GARCH effect}} \\
&+ \underbrace{\beta tr(\mathbf{P}\mathbf{L})\mathbf{x}'\mathbf{s}}_{\text{Linear dependence effect}} + \underbrace{\beta tr\left(\mathbf{P}\mathbf{Q}\right)\mathbf{x}'\mathbf{S}\mathbf{x}}_{\text{Quadratic dependence effect}} \\
&+ \underbrace{\beta\left[\mathbf{c}'\mathbf{L}\mathbf{c} + \mathbf{g}'\mathbf{L}\mathbf{g}\right]\mathbf{x}'\mathbf{s}}_{\text{Linear + ARCH/GARCH effect}} + \underbrace{\beta\left[\mathbf{c}'\mathbf{Q}\mathbf{c} + \mathbf{g}'\mathbf{Q}\mathbf{g}\right]\mathbf{x}'\mathbf{S}\mathbf{x}}_{\text{Quadratic + ARCH/GARCH effect}} \},
\end{aligned} \tag{5}$$

where

$$I = \beta tr\left(\mathbf{P}\mathbf{K}\right) - 2\beta\left(\mathbf{A}\mathbf{x} + \mathbf{B}\mathbf{i}\right)'\mathbf{p} + \beta\left(\mathbf{A}\mathbf{x} + \mathbf{B}\mathbf{i}\right)'\mathbf{P}\left(\mathbf{A}\mathbf{x} + \mathbf{B}\mathbf{i}\right).$$

The first line in the Bellman equation shows the terms found in a standard homoscedastic problem; the next two lines arise if there are ARCH or GARCH terms in the covariance structure; the fourth line shows the effect of state-dependent components and the last line shows the interaction if both of these effects are present. The first-order necessary condition for the minimum problem on the right side of equation (5) yields the optimal linear feedback

[4] We omit time subscripts in an equation whenever it includes only variables observed in the same period.

rule for the policy vector $\mathbf{i}$:

$$\mathbf{i} = \mathbf{f} - \mathbf{F}\mathbf{x} \tag{6}$$

$$\mathbf{f} = \left(\mathbf{W} + \beta\mathbf{B}'\mathbf{P}\mathbf{B}\right)^{-1}\left(\mathbf{W}\mathbf{i}^{*} + \mathbf{H}'\mathbf{x}^{*} + \beta\mathbf{B}'\mathbf{p}\right) \tag{7}$$

$$\mathbf{F}= \left(\mathbf{W} + \beta\mathbf{B}'\mathbf{P}\mathbf{B}\right)^{-1}\left(\mathbf{H}' + \beta\mathbf{B}'\mathbf{P}\mathbf{A}\right). \tag{8}$$

After substituting the trial solution (4) into the left side of equation (5) and the optimizer (6) - (8) into the right side, collection of the coefficients for the quadratic terms in $\mathbf{x}$ and those for $\mathbf{w}$ and $\boldsymbol{\Sigma}$ gives:

$$\mathbf{P} = \mathbf{R} + \beta\mathbf{A}'\mathbf{P}\mathbf{A} - \left(\mathbf{H} + \beta\mathbf{A}'\mathbf{P}\mathbf{B}\right)\left(\mathbf{W} + \beta\mathbf{B}'\mathbf{P}\mathbf{B}\right)^{-1}\left(\mathbf{H}' + \beta\mathbf{B}'\mathbf{P}\mathbf{A}\right) \tag{9}$$
$$+\beta tr\left(\mathbf{P}\mathbf{Q}\right)\mathbf{S} + \beta\mathbf{c}'\mathbf{Q}\mathbf{c}\mathbf{S} + \beta\mathbf{g}'\mathbf{Q}\mathbf{g}\mathbf{S}$$

$$\mathbf{c}\mathbf{c}' = \beta\mathbf{C}\left(\mathbf{P}+\mathbf{c}\mathbf{c}'+\mathbf{g}\mathbf{g}'\right)\mathbf{C}' \tag{10}$$

$$\mathbf{g}\mathbf{g}' = \beta\mathbf{G}(\mathbf{P} + \mathbf{c}\mathbf{c}' + \mathbf{g}\mathbf{g}')\mathbf{G}'. \tag{11}$$

Equating the coefficients for the linear terms gives

$$\mathbf{p} = \left[\mathbf{I}-\beta\left(\mathbf{A} - \mathbf{B}\mathbf{F}\right)'\right]^{-1}\left\{\begin{array}{c}\left(\mathbf{R} - \mathbf{F}'\mathbf{H}'\right)\mathbf{x}^{*} - \left(\mathbf{F}'\mathbf{W} - \mathbf{H}\right)\mathbf{i}^{*} \\ -\frac{1}{2}\beta\left[tr\left(\mathbf{P}\mathbf{L}\right) + \mathbf{c}'\mathbf{L}\mathbf{c} + \mathbf{g}'\mathbf{L}\mathbf{g}\right]\mathbf{s}\end{array}\right\}, \tag{12}$$

while collecting the constant terms gives

$$k = (1-\beta)^{-1}\left\{\left(\mathbf{f} - \mathbf{i}^{*}\right)'\mathbf{W}\left(\mathbf{f} - \mathbf{i}^{*}\right) + \mathbf{x}^{*\prime}\left[\mathbf{R}\mathbf{x}^{*} - 2\mathbf{H}\left(\mathbf{f} - \mathbf{i}^{*}\right)\right]\right. \tag{13}$$
$$\left. + \beta\mathbf{f}'\mathbf{B}'\left(\mathbf{P}\mathbf{B}\mathbf{f} - 2\mathbf{p}\right) + \beta[tr(\mathbf{P}\mathbf{K}) + \mathbf{c}'\mathbf{K}\mathbf{c} + \mathbf{g}'\mathbf{K}\mathbf{g}]\right.$$

Equation (9) is a matrix Riccati difference equation for the symmetric matrix $\mathbf{P}$, while (10) and (11) are discrete Lyapunov equations for the square matrices $\mathbf{g}\mathbf{g}'$ and $\mathbf{c}\mathbf{c}'$ respectively. This system can be solved by numerical iteration, starting from initial values for $\mathbf{P}$, $\mathbf{g}$ and $\mathbf{c}$. This is recursive: given the solution for $\mathbf{P}$ in (9) the solutions for $\mathbf{g}$ and $\mathbf{c}$ are obtained by joint numerical iteration of (10) and (11). Substituting $\mathbf{P}$, $\mathbf{g}$ and $\mathbf{c}$ into (12) and (13) then gives the solutions for $\mathbf{p}$ and k.

Equations (6) to (13) show the solution to the heteroscedastic optimal control problem. This encompasses a number of special cases that are now discussed separately to highlight how alternative specifications of the variance structure might alter the optimal feedback rule.[5]

2.2.1 Homoscedastic variance

If the variance structure is homoscedastic then equation (3) reduces to $\mathbf{\Sigma}_{t+1} = \mathbf{K}$; the trial solution is $V(\mathbf{x}) = k - 2\mathbf{x}'\mathbf{p} + \mathbf{x}'\mathbf{P}\mathbf{x}$; and the Bellman equation in (5) includes only the first line since the matrices $\mathbf{A}$, $\mathbf{G}$, $\mathbf{L}$ and $\mathbf{Q}$ are zero matrices and the vectors $\mathbf{g}$ and $\mathbf{c}$ are both null vectors. Differentiation with respect to the policy vector $\mathbf{i}$ yields the same optimal feedback rule as in equations (6) - (8), but with:

$$\hat{\mathbf{P}} = \mathbf{R} + \beta\mathbf{A}'\hat{\mathbf{P}}\mathbf{A} - \left(\mathbf{H} + \beta\mathbf{A}'\hat{\mathbf{P}}\mathbf{B}\right)\left(\mathbf{W} + \beta\mathbf{B}'\hat{\mathbf{P}}\mathbf{B}\right)^{-1}\left(\mathbf{H}' + \beta\mathbf{B}'\hat{\mathbf{P}}\mathbf{A}\right) \tag{14}$$

$$\hat{\mathbf{p}} = \left[\mathbf{I} - \beta\left(\mathbf{A} - \mathbf{B}\mathbf{F}\right)'\right]^{-1}\left[\left(\mathbf{R} - \mathbf{F}'\mathbf{H}'\right)\mathbf{x}^* - \left(\mathbf{F}'\mathbf{W} - \mathbf{H}\right)\mathbf{i}^*\right] \tag{15}$$

and

$$\hat{k} = \{(\mathbf{f} - \mathbf{i}^*)'\mathbf{W}(\mathbf{f} - \mathbf{i}^*) + \mathbf{x}^{*\prime}[\mathbf{R}\mathbf{x}^* - 2\mathbf{H}(\mathbf{f} - \mathbf{i}^*)] + \beta\mathbf{f}'\mathbf{B}'\left(\hat{\mathbf{P}}\mathbf{B}\mathbf{f} - 2\hat{\mathbf{p}}\right) + \beta tr\left(\hat{\mathbf{P}}\mathbf{K}\right)\}(1-\beta)^{-1}.$$

The solutions $\hat{\mathbf{P}}$ and $\hat{\mathbf{p}}$ are independent of the variance structure, therefore implying that the optimal feedback rule satisfies the CE principle.

2.2.2 GARCH variance

When the variance structure is driven by a GARCH process, equation (3) reduces to $\mathbf{\Sigma}_{t+1} = \mathbf{K} + \mathbf{A}'\mathbf{w}_t\mathbf{w}_t'\mathbf{A} + \mathbf{G}'\mathbf{\Sigma}_t\mathbf{G}$. The trial solution is given by (4), being different from that used for the homoscedastic case. The Bellman equation in (5) does not include the last two lines since the matrices $\mathbf{L}$ and $\mathbf{Q}$ are both equal to zero matrices. The solutions for $\mathbf{P}$ and $\mathbf{p}$ are still given by the equations (14) and (15). Thus the policy rule is as in the homoscedastic case.

[5] Appendix A provides more details on the computation of the solution in equations (6) to (13).

This result shows that under GARCH volatility the CE principle still holds, since the feedback rule in equations (6) - (8) is identical to the decision rule for the corresponding nonstochastic linear regulator problem.[6] Consequently the conditions sufficient for the stability of the solution to the homoscedastic optimal control problem are also sufficient for the stability of the solution with GARCH. This however affects welfare. Relative to the homoscedastic case, the value function (4) includes the non-negative terms $\mathbf{g}'\mathbf{w}\mathbf{w}'\mathbf{g}$ and $\mathbf{c}'\mathbf{\Sigma}\mathbf{c}$. In addition, the constant k changes under GARCH, as it includes the positive term $\beta\left(\mathbf{g}'\mathbf{K}\mathbf{g}+\mathbf{c}'\mathbf{K}\mathbf{c}\right)$ that increases the welfare loss because $\mathbf{K}$ is positive definite.

2.2.3 State dependence

The CE principle no longer holds if the variance structure includes linear and quadratic state-dependent components. Comparing (12) with (15) we can see that linear state dependence means that the last term in the square brackets on the right side of equation (12) is not zero. This shifts $\mathbf{p}$ and hence the intercepts $\mathbf{f}$ in the optimal feedback rule through (7). This shift is the sum of two effects. The first is direct, working through the term $\frac{1}{2}\beta tr\left(\mathbf{PL}\right)\mathbf{s}$ that occurs whenever there is linear state dependence in the variance structure. The second term $-\frac{1}{2}\beta\left(\mathbf{c}'\mathbf{L}\mathbf{c}+\mathbf{g}'\mathbf{L}\mathbf{g}\right)\mathbf{s}$ is a secondary effect that arises only when there is both GARCH and linear state dependence. Importantly, because $\mathbf{L}$ does not appear in the solution for $\mathbf{P}$ given by equations (9) - (11), linear state dependence does not affect the response coefficients $\mathbf{F}$ in (8). However, it affects the welfare loss from (4) as it changes $\mathbf{p}$ and consequently the constant term k in (13).

The impact of quadratic state dependence can be seen by comparing the solutions for $\mathbf{P}$ in equations (9) and (14). Under quadratic dependence, the term $\mathbf{Q}$ is no longer zero. This adds three extra terms on the right side of $\mathbf{P}$ in equations (9). The first, $\frac{1}{2}\beta tr\left(\mathbf{PQ}\right)\mathbf{S}$, is a direct effect that occurs whenever there is quadratic state dependence. The next two terms, $\beta\mathbf{c}'\mathbf{Q}\mathbf{c}\mathbf{S}$ and $\beta\mathbf{g}'\mathbf{Q}\mathbf{g}\mathbf{S}$, show a secondary effect on the response coefficients that occurs if there are both quadratic and GARCH components in the variance structure. Since $\mathbf{PQ}$ is positive semi-definite, and $\mathbf{c}'\mathbf{Q}\mathbf{c}$ and $\mathbf{g}'\mathbf{Q}\mathbf{g}$ are non-negative, these three terms are non-

[6] This result also applies to other time-varying specifications of the volatility that, like GARCH, are exogenous to the state variables, as for example the stochastic volatility and GARCH-X specification.

negative. They shift $\mathbf{P}$ through (9) and therefore $\mathbf{p}$ through (12). This affects the coefficients $\mathbf{f}$ and $\mathbf{F}$ in the optimal feedback rule through (7) and (8) and hence the welfare loss through (4).

In summary, there is a hierarchy of effects if the error structure is heteroscedastic. Quadratic state dependence affects all the coefficients in the optimal feedback rule as well as welfare. Linear state dependence affects the intercept of the optimal feedback rule and welfare. GARCH can reinforce these effects, but on its own this only affects welfare.

2.2.4 The Certainty-Equivalent Transform (CET)

Although the CE principle does not hold under linear-quadratic state-dependent volatility, a CET of the Bellman equation can be obtained by appropriately consolidating the linear and quadratic terms of the variance structure with the linear and quadratic terms in the objective function of the decision maker. This allows the value function to be expressed in the canonical homoscedastic form. Standard dynamic programming algorithms can then be used to solve problems that do not satisfy CE. We derive the CET for the general case of the heteroscedastic variance structure in equation (3).

The Bellman equation (5) can be re-parametrized in a certainty-equivalent form by setting $\mathbf{L}$ and $\mathbf{Q}$ to zero and replacing the welfare parameters $\mathbf{R}$, $\mathbf{x}^*$, $\mathbf{i}^*$ in (1) and k in (4) with $\widetilde{\mathbf{R}}$, $\widetilde{\mathbf{x}}^*$, $\widetilde{\mathbf{i}}^*$ and $\widetilde{k}$ to obtain the re-parametrized objective function:

$$\sum_{t=0}^{\infty}\beta^t E_t\left[\begin{array}{c}(\mathbf{x}_t-\widetilde{\mathbf{x}}^*)'\,\widetilde{\mathbf{R}}\,(\mathbf{x}_t-\widetilde{\mathbf{x}}^*)+\left(\mathbf{i}_t-\widetilde{\mathbf{i}}^*\right)'\mathbf{W}\left(\mathbf{i}_t-\widetilde{\mathbf{i}}^*\right)\\ +2\,(\mathbf{x}_t-\widetilde{\mathbf{x}}^*)'\,\mathbf{H}\left(\mathbf{i}_t-\widetilde{\mathbf{i}}^*\right)\end{array}\right], \tag{16}$$

where

$$\begin{aligned}
\widetilde{\mathbf{R}} &= \mathbf{R}+\beta\left[tr\left(\mathbf{PQ}\right)+\mathbf{c}'\mathbf{Qc}+\mathbf{g}'\mathbf{Qg}\right]\mathbf{S}\\
\widetilde{\mathbf{x}}^* &= [\widetilde{\mathbf{R}}-\mathbf{HW}^{-1}\mathbf{H}']^{-1}\{(\mathbf{R}-\mathbf{HW}^{-1}\mathbf{H}')\mathbf{x}^*-\frac{1}{2}\beta\mathbf{s}[tr(\mathbf{PL})+\mathbf{c}'\mathbf{Lc}+\mathbf{g}'\mathbf{Lg}]\}\\
\widetilde{\mathbf{i}}^* &= \mathbf{i}^*+\mathbf{W}^{-1}\mathbf{H}'(\mathbf{x}^*-\widetilde{\mathbf{x}}^*)\\
\widetilde{k} &= k+\beta^{-1}[\mathbf{x}^{*\prime}\mathbf{Rx}^*-\widetilde{\mathbf{x}}^{*\prime}\widetilde{\mathbf{R}}\widetilde{\mathbf{x}}^*-(\mathbf{i}^*-\widetilde{\mathbf{i}}^*)'\mathbf{W}(\mathbf{i}^*-\widetilde{\mathbf{i}}^*)-2\widetilde{\mathbf{x}}^{*\prime}\mathbf{HW}^{-1}\mathbf{H}'(\mathbf{x}^*-\widetilde{\mathbf{x}}^*)].
\end{aligned}$$

It then follows that the trial solution can be expressed as

$$\widetilde{V}\left(\mathbf{x},\mathbf{w},\boldsymbol{\Sigma}\right)=\widetilde{k}-2\mathbf{x}'\widetilde{\mathbf{p}}+\mathbf{x}'\widetilde{\mathbf{P}}\mathbf{x}+\mathbf{c}'\mathbf{w}\mathbf{w}'\mathbf{c}+\mathbf{g}'\boldsymbol{\Sigma}\mathbf{g} \tag{17}$$

and the certainty-equivalent Bellman equation can be written as:

$$\begin{aligned}\widetilde{V} = \min_{\mathbf{i}}[&(\mathbf{x}-\widetilde{\mathbf{x}}^{*})'\,\widetilde{\mathbf{R}}\,(\mathbf{x}-\widetilde{\mathbf{x}}^{*})+(\mathbf{i}-\widetilde{\mathbf{i}}^{*})'\mathbf{W}(\mathbf{i}-\widetilde{\mathbf{i}}^{*})+2\,(\mathbf{x}-\widetilde{\mathbf{x}}^{*})'\,\mathbf{H}(\mathbf{i}-\widetilde{\mathbf{i}}^{*}) \\ &+\beta\widetilde{k}+I+\beta tr((\widetilde{\mathbf{P}}\mathbf{C}'\mathbf{w}\mathbf{w}'\mathbf{C})+\beta tr(\widetilde{\mathbf{P}}\mathbf{G}'\boldsymbol{\Sigma}\mathbf{G}) \\ &+\beta\mathbf{c}'(\mathbf{K}+\mathbf{C}'\mathbf{w}\mathbf{w}'\mathbf{C}+\mathbf{G}'\boldsymbol{\Sigma}\mathbf{G})\mathbf{c}+\beta\mathbf{g}'(\mathbf{K}+\mathbf{C}'\mathbf{w}\mathbf{w}'\mathbf{C}+\mathbf{G}'\boldsymbol{\Sigma}\mathbf{G})\mathbf{g}],\end{aligned} \tag{18}$$

where I still defined as in section 2.2, with $\widetilde{\mathbf{P}}$ and $\widetilde{\mathbf{p}}$ replacing $\mathbf{P}$ and $\mathbf{p}$ respectively. Equation (18) shows that after transformation the Bellman equation has the same structure as in the GARCH model and thus satisfies the CE principle.

Appendix B derives the CET and shows that the Bellman equations in (5) and (18) are mathematically equivalent. Differentiation of (18) with respect to the vector of policy instruments $\mathbf{i}$ yields the feedback rule (6) but with coefficients determined as:

$$\widetilde{\mathbf{f}}=\left(\mathbf{W}+\beta\mathbf{B}'\mathbf{P}\mathbf{B}\right)^{-1}\left(\mathbf{W}\widetilde{\mathbf{i}}^{*}+\mathbf{H}'\widetilde{\mathbf{x}}^{*}+\beta\mathbf{B}'\widetilde{\mathbf{p}}\right) \tag{19}$$

$$\widetilde{\mathbf{F}}=\left(\mathbf{W}+\beta\mathbf{B}'\widetilde{\mathbf{P}}\mathbf{B}\right)^{-1}\left(\mathbf{H}'+\beta\mathbf{B}'\widetilde{\mathbf{P}}\mathbf{A}\right). \tag{20}$$

where:

$$\widetilde{\mathbf{P}}=\tilde{\mathbf{R}}+\beta\mathbf{A}'\widetilde{\mathbf{P}}\mathbf{A}-\left(\mathbf{H}+\beta\mathbf{A}'\widetilde{\mathbf{P}}\mathbf{B}\right)\left(\mathbf{W}+\beta\mathbf{B}'\widetilde{\mathbf{P}}\mathbf{B}\right)^{-1}\left(\mathbf{H}'+\beta\mathbf{B}'\widetilde{\mathbf{P}}\mathbf{A}\right), \tag{21}$$

$$\widetilde{\mathbf{p}}=\left[\mathbf{I}-\beta\left(\mathbf{A}-\mathbf{B}\mathbf{F}\right)'\right]^{-1}\left[\left(\tilde{\mathbf{R}}-\mathbf{F}'\mathbf{H}'\right)\widetilde{\mathbf{x}}^{*}-\left(\mathbf{F}'\mathbf{W}-\mathbf{H}\right)\widetilde{\mathbf{i}}^{*}\right]. \tag{22}$$

The solution to the transformed problem is equivalent to that in equations (6) to (13). The definition of $\widetilde{\mathbf{R}}$ implies that the solution for $\mathbf{P}$ from (9) is the same as the solution for $\widetilde{\mathbf{P}}$ from (21). Thus $\widetilde{\mathbf{F}}=\mathbf{F}$. The equality between $\mathbf{p}$ and $\widetilde{\mathbf{p}}$ is shown by replacing $\widetilde{\mathbf{x}}^{*}$ and $\widetilde{\mathbf{i}}^{*}$ in (22) and then simplifying. Substitution of $\widetilde{\mathbf{x}}^{*}$ and $\widetilde{\mathbf{i}}^{*}$ into (20) shows that $\tilde{\mathbf{f}}=\mathbf{f}$. Thus the policy rules from the general and the transformed control problem are the same. The value functions (4) and (17) differ by a constant, but this does not affect the decision rules.

The CET shows how heteroscedasticity affects the optimal policy. Since $\mathbf{Q}$ is non-negative definite, the scalar $[tr(\mathbf{PQ}) + \mathbf{c}'\mathbf{Qc} + \mathbf{g}'\mathbf{Qg}]$ in $\widetilde{\mathbf{R}}$ is non-negative and adds to the welfare weights of the variables identified by the selection vector $\mathbf{s}$ (and hence the matrix $\mathbf{S} = \mathbf{ss}'$). Similarly the non-negative scalar $[tr(\mathbf{PL}) + \mathbf{c}'\mathbf{Lc} + \mathbf{g}'\mathbf{Lg}]$ in $\widetilde{\mathbf{x}}^*$ and the selection vector $\mathbf{s}$ shift the intercept vector in (19).

The solution under the CET is stable since the matrix $\widetilde{\mathbf{R}}$ is positive semidefinite while the matrix $\mathbf{W}$ is unchanged in the transformed problem.

3 An illustrative empirical model

3.1 Data and model

We specify a second order VAR model of the U.S. economy based on 242 observations from 1953:1 to 2013:2 of unemployment (u_t), inflation (π_t) and the rate of interest (r_t).[7]

Let $\mathbf{z}_t = \begin{bmatrix} \mathbf{y}_t' & r_t \end{bmatrix}'$, with $\mathbf{y}_t = \begin{bmatrix} u_t & \pi_t \end{bmatrix}'$ denoting the block of non-policy variables, and $\mathbf{e}_t = \begin{bmatrix} \mathbf{e}_{\mathbf{y}t}' & e_{rt} \end{bmatrix}'$ denoting the corresponding vector of residuals. Then the VAR model can be written as:

$$\mathbf{z}_{t+1} = \mathbf{\Phi}_1 \mathbf{z}_t + \mathbf{\Phi}_2 \mathbf{z}_{t-1} + \mathbf{e}_t \tag{23}$$

with $\mathbf{\Phi}_1 = \begin{bmatrix} \mathbf{\Phi}_1^1 & \mathbf{\Phi}_1^2 \end{bmatrix}'$ and $\mathbf{\Phi}_2 = \begin{bmatrix} \mathbf{\Phi}_2^1 & \mathbf{\Phi}_2^2 \end{bmatrix}'$. The residuals can be written as $\mathbf{e}_{t+1} = \mathbf{\Xi}_{t+1}\mathbf{v}_{t+1}$, with $\mathbf{v}_t \sim N(\mathbf{0}, \mathbf{I})$, and their variance, $E[\mathbf{\Xi}_{t+1}\mathbf{\Xi}_{t+1}' | \mathbf{z}_t, \mathbf{e}_t, \mathbf{\Omega}_t]$, as:

$$\mathbf{\Omega}_{t+1} = \mathbf{\Omega_0} + \mathbf{\Omega_1} \mathbf{z}_t'\mathbf{s} + \mathbf{\Omega_2} \mathbf{z}_t'\mathbf{S}\mathbf{z}_t + \mathbf{M}\mathbf{e}_t\mathbf{e}_t'\mathbf{M}' + \mathbf{N}\mathbf{\Omega}_t\mathbf{N}' \tag{24}$$

where $\mathbf{\Omega}_0$, $\mathbf{\Omega}_1$ and $\mathbf{\Omega}_2$ are 3×3 real and symmetric matrices; $\mathbf{M}$ and $\mathbf{N}$ are 3×3 real and diagonal matrices; $\mathbf{s}' = \begin{bmatrix} 0 & 1 & 0 \end{bmatrix}$ so that $\mathbf{z}_\mathbf{t}'\mathbf{s} = \pi_t$; and $\mathbf{z}_t'\mathbf{S}\mathbf{z}_t = \pi_t^2$ (since $\mathbf{S} = \mathbf{ss}'$).

Equation (24) describes an encompassing (EN) model of the variance structure that nests four alternative models of volatility: the homoscedastic (HO) model ($\mathbf{\Omega}_t = \mathbf{\Omega}_0$); the linear-dependence (LN) model ($\mathbf{\Omega}_{t+1} = \mathbf{\Omega}_0 + \mathbf{\Omega}_1 \mathbf{z}_\mathbf{t}'\mathbf{s}$); the linear-quadratic (LQ) model

[7] We use the civilian unemployment rate for all workers over the age of 16; the annual percentage change in the implicit GDP price deflator; and the three-month Treasury bills. The data are demeaned, consistent with (2). The data are from http://research.stlouisfed.org/fred2.

($\mathbf{\Omega}_{t+1} = \mathbf{\Omega}_0 + \mathbf{\Omega}_1 \mathbf{z}_t' \mathbf{s} + \mathbf{\Omega}_2 \mathbf{z}_t' \mathbf{S} \mathbf{z}_t$); and the pure GARCH ($GH$) model ($\mathbf{\Omega}_{t+1} = \mathbf{\Omega_0} + \mathbf{M}\mathbf{e}_t\mathbf{e}_t'\mathbf{M}' + \mathbf{N}\mathbf{\Omega}_t\mathbf{N}'$).

The unconditional covariance matrix of the residuals $\mathbf{\Omega}$ is determined from the unconditional expectations on the right side of (24), as explained in Appendix C.1. Since $\mathbf{\Omega}$ is a real symmetric positive definite matrix, the triangular factorization $\mathbf{\Omega} = \mathbf{TDT}'$ applies, where $\mathbf{T}$ is a 3×3 lower triangular matrix with $\sigma_{\pi u}$, σ_{ru} and $\sigma_{r\pi}$ being the off-diagonal items and $\mathbf{D}$ is a 3×3 diagonal matrix with σ_{uu}, $\sigma_{\pi\pi}$ and σ_{rr} in the main diagonal.

3.2 Optimal monetary policy

The Fed chooses the sequence $\{r_t\}_{t=0}^{\infty}$ that minimizes the loss function

$$\sum_{t=0}^{\infty} \beta[\lambda_u u_t^2 + \lambda_\pi \pi_t^2 + \lambda_{\Delta r} (\Delta r_t)^2],$$

where u_t^2, π^2 and $(\Delta r)^2$ are the volatility of unemployment, inflation and changes in the policy instrument; while λ_u, λ_π and $\lambda_{\Delta r}$ are weights attached to each of the three goals respectively.[8] The optimization is subject to the constraints described by the non-policy block of the VAR in equations (23) and (24). As described in Appendix C.2, this is a special case of the general model in sections 2.1 and 2.2. Using these restrictions to simplify equation (16) and $\widetilde{\mathbf{R}}$, $\widetilde{\mathbf{x}}^*$, $\widetilde{\mathbf{i}}^*$ and $\widetilde{k}$ gives the CET of the loss function

$$\sum_{t=0}^{\infty} \beta \left\{ \lambda_u u_t^2 + \widetilde{\lambda}_\pi (\pi_t - \widetilde{\pi}^*) + \lambda_{\Delta r} \left[(r_t - \widetilde{r}^*)^2 + (r_{t-1} - \widetilde{r}^*)^2 - 2r_{t-1}r_t \right] \right\}$$

where

$$\begin{aligned}
\widetilde{\lambda}_\pi &= \lambda_\pi + \beta[tr\,(\mathbf{PQ}) + \mathbf{c}'\mathbf{Qc} + \mathbf{g}'\mathbf{Qg}] \geq \lambda_\pi \\
\widetilde{\pi}^* &= -\beta[tr\,(\mathbf{PL}) + \mathbf{c}'\mathbf{Lc} + \mathbf{g}'\mathbf{Lg}]/2\widetilde{\lambda}_\pi \\
\widetilde{r}^* &= \widetilde{\pi}^* \\
\widetilde{k} &= (1-\beta)^{-1} \beta tr(\mathbf{PK}) - \beta^{-1}\widetilde{\lambda}_\pi (\widetilde{\pi}^*)^2.
\end{aligned}$$

[8] Woodford (2003) shows that quadratic loss functions provide a good approximation to the expected lifetime utility of a representative household derived from a fully micro-founded macroeconomic model of the economy, in which inflation brings efficiency costs by distorting relative prices.

The re-parametrization, based on the solution in Appendix B, illustrates the effect of GARCH and state-dependent volatility on the optimal feedback rule. The term $\beta tr(\mathbf{PQ}) \geq 0$, due to the presence of quadratic dependence in (24), makes policy more aggressive as would an increase in the welfare weight λ_π. GARCH further reinforces this effect since both $\mathbf{c'Qc}$ and $\mathbf{g'Qg}$ are non-negative. The linear dependence term $\mathbf{L}$ reduces the effective target on inflation and interest rates from $\widetilde{r}^* = \pi^* = 0$ to $\widetilde{r}^* = \tilde{\pi}^* \leq 0$ since $\beta tr(\mathbf{PL}) \geq \mathbf{0}$. Provided that $\pi^* \geq 0$ then $\tilde{\pi}^* \leq \pi^*$ (since $\lambda_\pi \leq \tilde{\lambda}_\pi$). Thus linear dependence reduces the effective target and steady-state inflation and interest rates as a reduction in π^* would. This effect is reinforced by the presence of GARCH as both $\mathbf{c'Lc}$ and $\mathbf{g'Lg}$ are non-negative. Further the intercept in the value function shifts from k to $\widetilde{k}$ but, as noted, this does not affect the decision rule.

The optimal feedback rule can be combined with the non-policy block equations to study the dynamic of the VAR model under the optimal policy. We denote the VAR models under the optimal feedback rule as HO^*, LN^*, LQ^*, GH^*, EN^*.[9]

We write the steady-state solution to the policy rate equation as $\overline{r} = \bar{\kappa} + \bar{\phi}_u \overline{u} + \bar{\phi}_\pi \bar{\pi}$, with $\bar{\kappa}$, $\bar{\phi}_u$ and $\bar{\phi}_\pi$ denoting the long-run coefficients of the policy rule. In particular, $\bar{\kappa} = 0$ for models based on mean-adjusted data, but can be non-zero when the steady-state is shifted by the linear dependence effect under the optimal rule. With $\overline{r} = \bar{\pi} = \bar{\pi}^*$ and $\overline{u} = 0$ we have $\bar{\pi}^* = \bar{\kappa}/(1 - \bar{\phi}_\pi)$, where the denominator is negative under the Taylor principle requiring $\bar{\phi}_\pi > 1$, see Woodford (2003). This defines the stationary rate of inflation implicit in the long-run solution of the empirical and optimal feedback rule.

3.3 Maximum likelihood estimation

The VAR model in equations (23) and (24) is estimated by ML. First, we estimate the parameters in (24) while fixing the parameters $\mathbf{\Phi}_1$ and $\mathbf{\Phi}_2$ of the transition system (23) at their OLS values. This gives estimates for the homoscedastic model HO and four heteroscedastic models, labelled as LNX, LQX, GHX and ENX respectively. Next, we re-estimate all

[9] Appendix C.2 describes how this dynamic optimization problem is mapped into a form compatible with equations (1), (2) and (3) for the purpose of dynamic optimization. It also describes the derivation of the VAR under the optimal feedback rule.

the parameters in (23) and (24) simultaneously.[10]

Panels A and B in Table 1 show the likelihood statistics from the ML estimates. Under the log-likelihood ratio (LR) test, all restricted models are rejected at the 5 per cent significance level against the unrestricted ENX model. To guard against over-fitting, Table 1 also reports the difference in the Schwarz approximation to the Posterior Odds ratio (SCA) proposed by Canova (2007). Under this criterion only the HO model is rejected against the ENX model. Panel B shows that simultaneous estimation of all parameters in equations (23) and (24) produces a further, though modest, improvement in fit.

Panels C and D in Table 1 report the coefficients of the unconditional covariance matrix of the VAR innovations, $\mathbf{\Omega}$, implied by the ML estimates. The differences in the unconditional variances across models are important in understanding the welfare results reported in Section 3.5. Model LQX delivers the lowest unconditional variances for all three variables. This is particularly evident for the unconditional variance of (orthogonalized) interest rate shocks, σ_{rr}. Consequently, the unconditional variances are also low under the EN (ENX) model. The unconditional covariances are relatively stable across alternative specifications of the variance structure. In most cases, the GARCH model yields marginally lower estimates of the unconditional covariances than the other models.

3.4 The optimal policy rule

Table 2 reports the long-run coefficients of the interest rate equation and the rate of inflation described in section 3.2. The first column of numbers (headed 'Empirical') reports the estimates from the empirical policy rule. With the exception of model EN, these are by construction the same for all these models. The remaining columns (headed 'Optimal') report the coefficients for the long-run optimal feedback rule. We consider four different sets of welfare weights. The first gives equal weight to the three goal variables. The others show the effect of halving the weight on each goal variable.

Reading across Table 2 shows the effect of moving from the empirical to the optimal rule under different welfare specifications. Reading down this table shows the effect of different

[10] Appendix D describes the likelihood and the impulse response functions. The results from the ML estimates are available upon request.

models of volatility on the policy rule, to illustrate the theoretical findings of section 2.

We highlight the following results. First, optimization of model HO to get HO^* increases the response coefficients, with changes in the the optimal rule being consistent with the alternative specifications of the welfare weights. Second, the policy rule coefficients from model GHX^* are the same as in HO^*, since GARCH satisfies CE. Third, linear dependence introduces a positive intercept ($\bar{\kappa}$) into the interest rate equation of model LNX^*, thereby reducing steady-state inflation ($\bar{\pi}^*$), while leaving the optimal response coefficients unchanged relative to models HO^* and GHX^*. Fourth, in the LQX^* model, quadratic dependence also makes policy more responsive to inflation. The shift in $\bar{\kappa}$ is slightly larger than in LNX^*, leading to lower steady-state inflation. Fifth, the coefficients from model ENX^* are very similar to those in model LQX^*, and the intercepts are only marginally higher. Thus the secondary effects of including GARCH as well as state-dependent volatility appear to be empirically small.[11] Six, comparison of the results for ENX and EN shows that re-estimating the transmission coefficients has little effect on the long-run policy rule.

3.5 Welfare analysis

Table 3 shows how heteroscedasticity affects the measurement of welfare by reporting the losses obtained from the stochastic simulation of the six models.[12] The analysis is presented along three different dimensions. The first column of numbers illustrates how heteroscedasticity affects welfare under the empirical rule. The third column of numbers shows the effects under the optimal policy rule. The last column shows the welfare gains from the optimization of policy. In parenthesis we report the welfare changes due to different variance structures relative to the homoscedastic model.

The first column shows that GARCH (GHX) and linear dependence (LNX) increase the loss, while the quadratic effect reduces it. The lowest loss occurs for model LQX, consistent with the observation in section 3.3 that this is less sensitive to interest rate shocks than other models. The comparison of models HO^* and GHX^* suggests that adding GARCH

[11] This is consistent with the result in Table 1 that changes in the variance structure are mainly driven by state-dependence rather than the GARCH.

[12] Appendix E describes the methodology used for the stochastic simulation.

to model HO^* is broadly the same as adding $GARCH$ to model HO.[13] There is nothing policy can do about the increase in variability, the variance of one goal variable can only be traded off against that of another. In LNX^* however, policy can reduce the overall volatility of the system by reducing the steady-state rate of inflation (as shown in Table 2). This mitigates the effect of introducing linear state-dependent heteroscedasticity.[14] Introducing quadratic state dependence into the empirical model lowers the welfare loss (about 4% on average across the four sets of welfare weights), but this reduction is greater (about 13% on average) under the optimal rule. It achieves this by combining a reduction in steady-state inflation with a more aggressive policy stance.

The last column shows that the gain from optimization is generally higher when allowing for inflation-conditional volatility in the variance structure. In model LNX^* it is optimal to lower the steady-state inflation rate, thereby shifting the trade-off and reducing the overall volatility of the system. This makes the optimization gain bigger than in the standard homoscedastic model, where the gain only reflects increase in the level of aggression in the systematic response of policy.[15] The LQX^* model, which combines a shift in the steady state with a more aggressive stance, gives an average welfare gain across the four specifications of preferences that is almost twice that implied by model HO^*. The welfare gains from the optimization of model GHX are broadly the same as under model HO, since both satisfy the CE principle. Models ENX and EN lead to an average welfare improvement, between 30-45 per cent, that is still higher than that from the optimization of the HO and GHX models.

The numerical analysis is based on a reduced-form model. In this respect, we followed a large literature on the implications of changes in monetary policy for macroeconomic dynamics and welfare also based on reduced-form models like ours. Examples include

[13] For example, the top panel of Table 3 shows that with equal welfare weights, this increases the loss by 8.43% compared with 8.55% with the empirical rule. Introducing GARCH into LQX^* to get ENX^* also has a similar effect under the optimal rule (an increase in the loss of $100\ln(5.9/5.18) = 13\%$) as it does under the empirical rule ($100\ln(7.30/6.34) = 14.1\%$). These increases are also similar under the alternative welfare specifications shown in the other panels.

[14] Looking at the top panel again, this increases the welfare loss by 8.97% comparing HO and LNX but by just 6.55% comparing HO^* and LNX^*, with similar reductions in the other panels.

[15] The average welfare gain across the four set of welfare specifications is about 11 per cent in model HO^* (compared to HO) and 17 per cent in model LNX^* (compared to LNX).

Bernanke and Mihov (1998), Sack (2000), Sims and Zha (2006a, 2006b) and Polito and Wickens (2012).

Reduced-form models are subject to the Lucas (1976) critique that the transition mechanism in the economy is in theory not invariant to policy changes.[16] The empirical relevance of this observation is however still debated. Our VAR model is based on Primiceri (2005) who finds no significant changes in the responses of inflation and unemployment to the policy rate under Burns, Volcker and Greenspan chairmanships of the Fed. Sims and Zha (2006a, 2006b) come to a similar conclusion. Polito and Wickens (2012) also find not significant evidence of the sort of structural instability predicted by the Lucas critique in a VAR of the US economy over the period 1964:2-2009:3. But to set against this evidence, see Benati and Surico (2009) and Benati (2010).

An alternative is to use a structural model like that of Rudebusch and Svensson (1999) that is less vulnerable to the Lucas critique. However, structural models employ dynamic restrictions which could induce heteroscedasticity through misspecification. Nevertheless, companion papers for the U.S., Polito and Spencer (2011a), and for the U.K., Polito and Spencer (2011b), give results based on Rudebusch and Svensson (1999) that are similar to those reported here, thus suggesting that these findings are likely to be robust across a wider range of models.

4 Conclusion

It is well known that the optimal feedback rule satisfies CE if volatility is homoscedastic. We show that this result also holds in any model in which the source of change in the variance structure is exogenous. However, volatility can also be state dependent. The CE principle no longer holds if changes in the variance structure are endogenously driven, because volatility affects welfare and state dependence puts the policy maker in a position to influence this. We show that if the variance structure is linear-quadratic, a CET of the optimal linear regulator problem can be obtained so that it resembles a standard homoscedastic model

[16] In Table 3, the Lucas critique does not apply to the welfare effects of heteroscedasticity under the empirical rule (first column of numbers) and the optimal rule for homoscedastic, GARCH and linear-dependent variance (third column of numbers) since these share the same policy response coefficients.

problem. This allows the researcher to use the algorithms and insights provided by existing methodologies.

Optimization under state-dependent volatility brings two main effects. Linear state dependence affects the overall variance of the system by shifting its steady state, while quadratic dependence changes the systematic component of policy. These two effects are mathematically equivalent to changes in the targets and welfare weights in the homoscedastic linear regulator problem. If GARCH is also present, this has the effect of amplifying these shifts.

5 Acknowledgments

We are grateful for comments and feedback from the editor, Fabio Canova, and three anonymous referees. We are also grateful for helpful suggestions to Christos Ioannidis, Gary Koop, James Morley, Ian Tonks and Mike Wickens. We thank participants at seminar presentations at the Bank of England, the European Central Bank, the Reserve Bank of Australia, Macquarie University, the University of Technology Sidney, University of New South Wales and the Universities of Bath, Cardiff, Strathclyde and York.

6 References

Ball L. 1992. Why Does Inflation Raise Inflation Uncertainty. *Journal of Monetary Economics* **29**: 371-388. DOI:10.1016/0304-3932(92)90032-W

Bekaert G, Cho S, Moreno A. 2010. New Keynesian Macroeconomics and the Term Structure. *Journal of Money, Credit, and Banking* **42**: 33-62. DOI: 10.1111/j.1538-4616.2009.00277.x

Benati L. 2010. Are policy counterfactual based on structural VAR's reliable? European Central Bank, Working Paper Series 1188.

Benati L, Surico P. 2009. VAR Analysis and the Great Moderation. *American Economic Review* **99**: 1636-1652. DOI: 10.1257/aer.99.4.1636

Bernanke B, Mihov I. 1998. Measuring Monetary Policy. *The Quarterly Journal of Economics* **113**: 869-902. DOI:10.1162/003355398555775

Blanchard O, Kahn C. 1980. The Solution of Linear Difference Models under Rational Expectations. *Econometrica* **48**: 1305-1311. DOI: 10.2307/1912186

Blanchard O, Dell'Ariccia G, Mauro P. 2010. Rethinking macroeconomic policy. SPN/10/03, IMF Staff Position Note.

Campbell JY, Pflueger C, Viceira LM. 2014. Monetary Policy Drivers of Bonds and Equity Risks. NBER Working Paper No. 20070.

Canova F. 2007. *Methods for Applied Macroeconomics Research*. Princeton University Press.

Caporale GM., Kontonikas A. 2009. The Euro and Inflation Uncertainty in the European Monetary Union. *Journal of International Money and Finance* **28**: 954-971. DOI:10.1016/ j.jimonfin.2008.09.004

Cox J, Ingersoll J, Ross S. 1985. A Theory of the Term Structure of Interest Rates. *Econometrica* **52**: 385-407. DOI: 10.2307/1911242

Courtadon G. 1982. The Pricing of Options on Default-Free Bonds. *The Journal of Financial and Quantitative Analysis* **17**: 75-100.

Dennis R. 2006. The Policy Preferences of the US Federal Reserve. *Journal of Applied Econometrics* **21**: 55-77. DOI: 10.1002/jae.808

Dothan U. 1978. On the Term Structure of Interest Rates. *Journal of Financial Economics* **6**: 59-69. DOI:10.1016/0304-405X(78)90020-X

Engle RF. 1982. Autoregressive Conditional Heteroskedasticity with Estimates of the Variance of the United Kingdom Inflation. *Econometrica* **50**: 907-1008. DOI: 10.2307/1912773

Engle RF, Kroner KF. 1995. Multivariate simultaneous generalized ARCH. *Econometric Theory.* **11**: 122-150. DOI: http://dx.doi.org/10.1017/S0266466600009063

Fountas S., Karanasos M, Kim J. 2002. Inflation and Output Growth Uncertainty and their Relationship with Inflation and Output Growth. *Economic Letters* **75**: 293-301. DOI:10.1016/S0165-1765(02)00009-5

Friedman M. 1977. Nobel Lecture: Inflation and Unemployment. *Journal of Political Economy* **85**: 451-472.

Hansen LP, Sargent TJ. 2008. *Robustness*. Princeton University Press.

Lucas R. 1976. Econometric Policy Evaluation: A Critique. In *The Phillips Curve* *and Labor Markets,* Brunner K, Meltzer A (eds). Carnegie-Rochester Conference Series on Public Policy 1. New York: American Elsevier, 19-46.

Ljunqvist L, Sargent TJ. 2004. *Recursive Macroeconomic Theory*, MIT: Cambridge, MA.

Okun A 1971. The Mirage of Steady Inflation. *Brookings Papers on Economic Activity* **2**: 485-498. DOI: 10.2307/2534234

Polito V, Spencer P. 2011a. Optimal Monetary Policy with Inflation Conditional Volatility. Mimeo available at www.finance.uts.edu.au/research/seminars/110307.pdf.

Polito V, Spencer P. 2011b. UK Macroeconomic Volatility and the Welfare Costs of Inflation. University of York. Discussion Paper 11/21.

Polito V, Wickens MR. 2012. Optimal Monetary Policy Using an Unrestricted VAR. *Journal of Applied Econometrics* **27**: 525-553. DOI: 10.1002/jae.1219

Primiceri G. 2005. Time-Varying Structural Vector Autoregressions and Monetary Policy. *Review of Economic Studies* **72**: 821-852. DOI: 10.1111/j.1467-937X.2005.00353.x

Rudebusch GD, Svensson L. 1999. Policy Rules for Inflation Targeting. In *Monetary Policy Rules*, Taylor J (eds) 203-262. Cambridge, Massachusetts: NBER.

Sack BP. 2000. Does the Fed Act Gradually? A VAR Analysis. *Journal of Monetary Economics* **46**: 229-256. DOI:10.1016/S0304-3932(00)00019-2

Sims C, Zha T. 2006a. Does Monetary Policy Generate Recessions? *Macroeconomic Dynamics* **10**: 231-272. DOI: http://dx.doi.org/10.1017/S136510050605019X

Sims C, Zha T. 2006b. Were There Regime Switches in U.S. Monetary Policy? *American Economic Review* **96**: 54-81. DOI: 10.1257/000282806776157678

Spencer PD. 2008. Stochastic Volatility in a Macro-Finance Model of the U.S. Term Structure of Interest Rates 1961-2004. *Journal of Money, Credit and Banking* **40**: 1177-1215. DOI: 10.1111/j.1538-4616.2008.00153.x

Woodford M. 2003. *Interest and Prices: Foundations of a Theory of Monetary Policy.* Princeton University Press.

Table 1: ML results

Likelihood and test statistics					
Panel A: First stage					
	HO	LNX	LQX	GHX	ENX
LogL	311.20	356.70	379.70	354.40	394.00
BIC	582.92	654.19	680.45	629.85	630.10
N	6	9	12	12	24
LR test	165.60	74.60	28.60	79.20	
SCA test	23.59	-12.04	-25.17	0.12	
M (χ^2_M)	18 (28.87)	15 (25)	12 (21.03)	12 (21.03)	
Panel B: Second stage					
	HO	LN	LQ	GH	EN
LogL	311.20	359.20	397.00	370.10	401.00
BIC	464.50	540.76	596.62	542.82	565.15
N	24	27	30	30	36
LR test	179.60	83.60	8.00	61.80	
SCA test	-50.32	-12.19	15.73	-11.16	
M (χ^2_M)	12 (21.03)	9 (16.92)	6 (12.59)	6 (12.59)	
ML estimates of the unconditional covariance matrix					
Panel C: First stage of ML estimation					
	HO	LNX	LQX	GHX	ENX
σ_{uu}	0.074	0.085	0.060	0.083	0.070
$\sigma_{\pi\pi}$	0.104	0.113	0.100	0.110	0.115
σ_{rr}	0.467	0.437	0.176	0.541	0.253
$\sigma_{u\pi}$	-0.037	-0.055	-0.040	-0.079	-0.048
σ_{ur}	-0.577	-0.559	-0.560	-0.665	-0.605
$\sigma_{r\pi}$	0.331	0.356	0.375	0.301	0.367
Panel D: Second stage of ML estimation					
		LN	LQ	GH	EN
σ_{uu}		0.091	0.064	0.082	0.061
$\sigma_{\pi\pi}$		0.120	0.105	0.110	0.106
σ_{rr}		0.441	0.234	0.545	0.275
$\sigma_{u\pi}$		-0.067	-0.091	-0.064	-0.085
σ_{ur}		-0.562	-0.576	-0.655	-0.564
$\sigma_{r\pi}$		0.342	0.358	0.299	0.348

Note: LogL is the log-likelihood; BIC is the Bayesian information criterion; N is the number of estimated parameters; LR test is the log-likelihood ratio test relative to $EN(X)$; SCA test is the Schwarz statistic. The LR and SCA tests have a χ^2_M distribution, with M being the number of restrictions based on the difference in the parameters of the unrestricted and restricted model. The figures in the χ^2_M rows are the 95% critical values. The ML estimates of the unconditional covariance matrix are based on the solution of the matrix equation (40) in the Online appendix.

Table 2: Long-run responses of estimated and optimal policy rules

	Empirical	Optimal weights			
		$\lambda_u = \lambda_\pi = \lambda_{\Delta r} = 1$	$\lambda_u = 0.5$; $\lambda_\pi = \lambda_{\Delta r} = 1$	$\lambda_\pi = 0.5$; $\lambda_u = \lambda_{\Delta r} = 1$	$\lambda_{\Delta r} = 0.5$; $\lambda_u = \lambda_\pi = 1$
	HO	HO^* ($\bar{\kappa} = \bar{\pi}^* = 0$)			
$\bar{\phi}_u$	-0.39	-1.97	-1.51	-2.04	-2.56
$\bar{\phi}_\pi$	1.38	4.00	4.38	2.76	5.06
	GHX	GHX^* ($\bar{\kappa} = \bar{\pi}^* = 0$)			
$\bar{\phi}_u$	-0.39	-1.97	-1.51	-2.04	-2.56
$\bar{\phi}_\pi$	1.38	4.00	4.38	2.76	5.06
	LNX	LNX^*			
$\bar{\kappa}$	(-)	0.61	0.53	0.43	1.09
$\bar{\phi}_u$	-0.39	-1.97	-1.51	-2.04	-2.56
$\bar{\phi}_\pi$	1.38	4.00	4.38	2.76	5.06
$\bar{\pi}^*$	(-)	-0.20	-0.16	-0.25	-0.27
	LQX	LQX^*			
$\bar{\kappa}$	(-)	1.04	0.93	0.71	1.90
$\bar{\phi}_u$	-0.39	-1.96	-1.52	-2.01	-2.55
$\bar{\phi}_\pi$	1.38	4.33	4.61	3.17	5.43
$\bar{\pi}^*$	(-)	-0.31	-0.26	-0.33	-0.43
	ENX	$ENX*$			
$\bar{\kappa}$	(-)	1.14	0.98	0.83	2.05
$\bar{\phi}_u$	-0.39	-1.96	-1.52	-2.01	-2.55
$\bar{\phi}_\pi$	1.38	4.29	4.57	3.14	5.38
$\bar{\pi}^*$	(-)	-0.35	-0.27	-0.39	-0.47
	EN	EN^*			
$\bar{\kappa}$	(-)	1.76	1.29	1.23	4.57
$\bar{\phi}_u$	-0.35	-1.20	-0.95	-1.32	-1.45
$\bar{\phi}_\pi$	1.19	3.42	3.51	2.73	4.17
$\bar{\pi}^*$	(-)	-0.73	-0.51	-0.71	-1.44

Note: the long-run interest rate rule is $\overline{r} = \bar{\kappa} + \bar{\phi}_u\overline{u} + \bar{\phi}_\pi\overline{\pi}$. The intercept is zero for all models under the empirical rule since the data is de-meaned prior to estimation and optimization, and also for models HO^* and GHX^* since these are certainty equivalent. Linear-dependence in the variance structure has the effect of inducing a positive $\bar{\kappa}$ intercept in the optimal feedback rule; this reduces the steady-state inflation rate by $\bar{\pi}^* = \bar{\kappa}/(1 - \bar{\phi}_\pi)$ in models LNX^*, LQX^*, ENX^* and EN. The loss function is described in section 3.2; λ_u, λ_π and $\lambda_{\Delta r}$ are weights attached to unemployment, inflation and changes in the rate of interest volatility respectively.

Table 3: Welfare measurement and heteroscedasticity

Empirical rule			**Optimal rule**			
Model	Welfare loss	change on HO (%)	Model/ Case	Welfare loss	change on HO^* (%)	Gain (%) from model Optimisation
Case 1:	$\lambda_\pi = \lambda_{\Delta r} = 1$					
HO	6.61		HO^*	5.91		11.19
GHX	7.2	(8.55)	GHX^*	6.43	(8.43)	11.31
LNX	7.23	(8.97)	LNX^*	6.31	(6.55)	13.61
LQX	6.34	(-4.17)	LQX^*	5.18	(-13.18)	20.21
ENX	7.3	(9.93)	ENX^*	5.9	(-0.17)	21.29
EN	6.46	(-2.30)	EN^*	4.66	(-23.76)	32.66
Case 2:	$\lambda_u = 0.5$; $\lambda_\pi = \lambda_{\Delta r} = 1$					
HO	5.53		HO^*	4.8		14.16
GHX	6.03	(8.66)	GHX^*	5.23	(8.58)	14.23
LNX	6.02	(8.49)	LNX^*	5.13	(6.65)	16.00
LQX	5.29	(-4.44)	LQX^*	4.21	(-13.12)	22.84
ENX	6.07	(9.32)	ENX^*	4.82	(0.42)	23.06
EN	5.12	(-7.70)	EN^*	3.67	(-26.84)	33.30
Case 3:	$\lambda_\pi = 0.5$; $\lambda_u = \lambda_{\Delta r} = 1$					
HO	4.69		HO^*	4.44		5.48
GHX	5.13	(8.97)	GHX^*	4.85	(8.83)	5.61
LNX	5.12	(8.77)	LNX^*	4.72	(6.12)	8.13
LQX	4.48	(-4.58)	LQX^*	3.95	(-11.69)	12.59
ENX	5.17	(9.74)	ENX^*	4.47	(0.67)	14.55
EN	4.82	(2.73)	EN^*	3.8	(-15.57)	23.78
Case 4:	$\lambda_{\Delta r} = 0.5$; $\lambda_\pi = \lambda_u = 1$					
HO	6.3		HO^*	5.26		18.04
GHX	6.84	(8.22)	GHX^*	5.7	(8.03)	18.23
LNX	6.93	(9.53)	LNX^*	5.6	(6.26)	21.31
LQX	6.08	(-3.55)	LQX^*	4.52	(-15.16)	21.65
ENX	7.01	(10.68)	ENX^*	5.15	(-2.11)	30.83
EN	6.18	(-1.92)	EN^*	3.96	(-28.39)	44.51

Note: This table shows the welfare losses obtained by simulating the various models under the empirical and optimal feedback rules. Reading down the table shows the effect of introducing different models of the variance structure upon the welfare rule. Reading across this table shows the effect of moving from the empirical interest rate rule to an optimal rule under different welfare specifications. The final column shows the logarithmic percentage increase in welfare.